Model: $Output = \beta_0 + \beta_1 \cdot Cost + \varepsilon$ (DV) (IV)

$N=90$
$K=2$ (β_0, β_1)

```
. regress output cost

      Source |       SS       df       MS              Number of obs =      90
-------------+------------------------------           F(  1,    88) =  880.73
       Model |  107.123089     1  107.123089           Prob > F      =  0.0000
    Residual |  10.7034329    88  .12162992            R-squared     =  0.9092
-------------+------------------------------           Adj R-squared =  0.9081
       Total |  117.826522    89  1.32389351           Root MSE      =  .34875

------------------------------------------------------------------------------
      output |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        cost |   .9691952    .032658    29.68   0.000     .9042942    1.034096
       _cons |  -14.12819   .4380397   -32.25   0.000    -14.99871   -13.25768
------------------------------------------------------------------------------
```

Model = SSM, Residual = SSE, Total = SST; MS Model = MSM, MS Residual = MSE; df: Model = $K-1$, Residual = $N-K$, Total = $N-1$

$F = \frac{MSM}{MSE}$, $R^2 = \frac{SSM}{SST}$, Adj $R^2 = 1-((N-1)/(N-K))(1-R^2)$

Root MSE $= \sqrt{MSE} = \sqrt{.1216} = S_\varepsilon$, $MSE = S_\varepsilon^2 = \hat{\sigma}_\varepsilon^2$

cost coef $= \frac{SP_{xy}}{SS_{xx}}$, _cons coef $= \bar{y} - b\bar{x}$; $t = b_1/SE_{b_1}$, $t = b_0/SE_{b_0}$

$SE_{b_1} = \sqrt{\frac{MSE}{SS_{xx}}} = \sqrt{\frac{S_\varepsilon^2}{SS_{xx}}}$

$SE_{b_0} = \sqrt{MSE\left(\frac{1}{n}+\frac{\bar{x}^2}{SS_{xx}}\right)} = \sqrt{S_\varepsilon^2\left(\frac{1}{n}+\frac{\bar{x}^2}{SS_{xx}}\right)}$

```
. tabstat output cost, stat(n mean)

   stats |    output      cost
---------+--------------------
       N |        90        90
    mean | -1.174309  13.36561
------------------------------
```

mean output = $\bar{y}$, mean cost = $\bar{x}$

```
. gen SSxx=(cost-13.36561)^2
. gen SSyy=(output-(-1.174309))^2
. gen SSxy=(cost-13.36561)*(output-(-1.174309))

. tabstat cost output SSxx SSyy SPxy, stat(mean sum)

   stats |      cost    output      SSxx      SSyy      SPxy
---------+--------------------------------------------------
    mean |  13.36561 -1.174309  1.267121  1.309184  1.228088
     sum |  1202.905 -105.6878  114.0409  117.8265  110.5279
------------------------------------------------------------
```

cost = X, output = Y; sum of SSyy = SST

```
. di 110.5279/114.0409
.96919526

. di -1.174309-.96919526*13.36561
-14.128195

. gen yhat=-14.12819+.9691952*cost
. gen ee=(output-yhat)^2
. tabstat ee, stat(sum)

    variable |       sum
-------------+----------
          ee |  10.70343
------------------------

. di 10.70343/(90-2)
.12162989

. di sqrt(.12162989/114.0409)
.03265802

. di .9691952/.03265802
29.677096

. di sqrt(.12162989*(1/90+13.36561^2/114.0409))
.43803968

. di -14.12819/.43803968
-32.25322

. di 117.826522-10.70343
107.12309
```

$\frac{SP_{xy}}{SS_{xx}} \rightarrow b_1$

$\bar{y} - b_1 \cdot \bar{x} \rightarrow b_0$

$\hat{y} = a + b \cdot x$

$e^2 = (y-\hat{y})^2$

$\sum \varepsilon^2 = \sum (y-\hat{y})^2 = SSE$

$MSE = \frac{SSE}{N-K}$

$SE_{b_1} = \sqrt{Var(b_1)} = \frac{MSE}{SS_{xx}}$

$t_{b_1} = \frac{b_1}{SE_{b_1}}$

$SE_{b_0} = \sqrt{Var(b_0)} = \sqrt{MSE\left(\frac{1}{n}+\frac{\bar{x}^2}{SS_{xx}}\right)}$

$t_{b_0} = \frac{b_0}{SE_{b_0}}$

$SSM = SST - SSE$

Regression

[DataSet5]

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	cost[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: output

$SSE = SST - SSM$

$MSE = SSE/(N-k)$ to be used in $\frac{MSM}{MSE}$, Var(b), Var(a)

$SEE = \sqrt{MSE}$ ← standard error of residual $S_\varepsilon = \sqrt{S_\varepsilon^2}$

MSE = Variance of residual $S_\varepsilon^2 = \sigma_\varepsilon^2$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.953[a]	.909	.908	.34875

a. Predictors: (Constant), cost

R Square: $\frac{SSM}{SST}$

Adjusted R Square: $1-[(N-1)/(N-k)](1-R^2)$

Std. Error of the Estimate: $\sqrt{MSE} = SEE$ standard error of residual (error)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig. (p-value)
1	Regression	SSM 107.123	K-1 1	MSM 107.123	880.730	.000[a]
	Residual	SSE 10.703	N-K 88	MSE .122	($\frac{MSM}{MSE}$)	
	Total	SST 117.827	N-1 89			

a. Predictors: (Constant), cost IV

b. Dependent Variable: output DV

p-value < .05, Reject H_0 of $\beta_0 = \beta_1 = 0$

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig. (p-value)
1	(Constant)	-14.128	.438		-32.253	.000
	cost	.969	.033	.953	29.677	.000

a. Dependent Variable: output

No!

$a = \bar{y} - b\bar{x}$

$b = \frac{SP_{xy}}{SS_{xx}}$

$SE_{b_1} = \sqrt{\frac{MSE}{SS_{xx}}}$

$SE_{b_0} = \sqrt{MSE\left(\frac{1}{n} + \frac{\bar{x}^2}{SS_{xx}}\right)} = \sqrt{S_\varepsilon^2\left(\frac{1}{n} + \frac{\bar{x}^2}{SS_{xx}}\right)}$

$\frac{b_1}{SE_{b_1}} = \frac{.969}{.033}$

$\frac{b_0}{SE_{b_0}} = \frac{-14.128}{.438}$

$b_1 \frac{S_x}{S_y} = .969 * \frac{1.13199}{1.150606}$

"For a standard deviation increase in x, y is expected to change by $b_1 \frac{S_x}{S_y}$ standard deviations, holding all other variabls constant."

★ This is not β_1 but "standardized b_1."